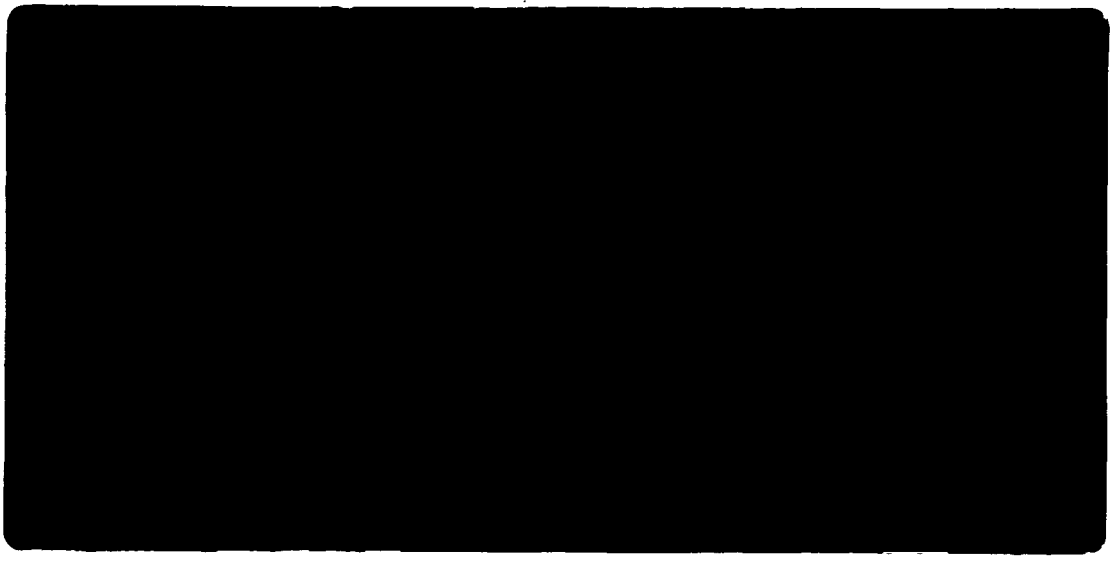
Institute of Paper Science and Technology
Atlanta, Georgia

IPST TECHNICAL PAPER SERIES



NUMBER 456

COMPATIBILITY OF FUNCTIONAL GROUPS IN K_{OW}-BASED QSARS

S. BANERJEE AND C.L. WILLIAMS

NOVEMBER 1992

Compatibility of Functional Groups in K_{OW}-Based QSARs

S. Banerjee and C.L. Williams

Submitted to
Environ. Toxicol. Chem.

Abstract

It is shown that nitro-compounds are particularly difficult to handle in simple K_{ow}-based QSARs owing to differences in their lipid phase activity coefficients. These differences can be corrected through inclusion of a term in octanol solubility. A procedure is developed for anticipating activity coefficient related deviations from the nature of the groups used.

Quantitative structure-activity relationships (QSARs) used in environmental work are most often based on K_{ow}, the octanol: water partition coefficient. A simple K_{ow}-based QSAR assumes that the activity coefficients of members of the series remain relatively constant in both the biological medium and in octanol. This is frequently the case across a homologous series, and good QSARs are most often found across structurally related compounds.

The activity coefficient of a solute in both octanol (γ_{oct}) and lipid varies with molecular size and with the nature of the solute-solvent interaction. The interaction term is usually the more important for mid-molecular weight compounds. Consider the QSAR of Deneer et al. (1) on the acute toxicity of nitro-aromatics toward the guppy, which is illustrated in Fig. 1. The correlation is poor. Much of the scatter reflects physico-chemical rather than biological effects since the relationship improves appreciably if a term in octanol solubility S_{oct} (determined as in ref. 2) is included. The resulting equation

$$\log LC_{50} = 3.82 - 0.52 \log K_{ow} + 0.93 \log S_{oct} \qquad (1)$$

is illustrated in Fig. 2. Similarly, Purdy (3) observed considerable scatter in a QSAR on the acute toxicity of mono-, di- and tri-substituted nitrobenzenes toward the fathead minnow

(Fig. 3). The line in Fig. 3 is for non-nitro compounds. Inclusion of an S_{oct} term leads to

$$\log LC_{50} = -1.49 - 0.64 \log K_{ow} - 1.09 \log S_{oct} \quad (2)$$

which corresponds to the much improved relationship illustrated in Fig. 4. It is noteworthy that while the coefficients of the K_{ow} term are similar in eq. 1 and 2, those for the S_{oct} term are of opposite sign. In eq. 1, the S_{oct} term depresses toxicity; in eq. 2, it enhances it. Thus, octanol solubility improves the QSARs for apparently different reasons. An analogous improvement has been previously observed in bioconcentration relationships (2).

S_{oct} is most likely to be constant over a range of solutes if the groups involved all have comparable interactions with octanol. One measure of nonideality that can be applied at the group level is the UNIFAC residual activity coefficient γ_r (4). This term reflects the enthalpic interaction of a group with its environment. Values of γ_r were calculated for the groups listed in Table 1 at infinite dilution in octanol at 25°C. Calculations for multiple groups or combinations of groups are easily made through the UNIFAC equation (4). In this paper we postulate that groups of similar γ_r in Table 1 will tend to be the most compatible in a simple K_{ow}-based QSAR, at least from a purely physicochemical point of view.

The nitro group is notable in that it has one of the highest γ_rs in Table 1. This suggests that for a simple K_{ow}-based QSAR to be successful for nitro compounds, each member of the series should contain the same number of nitro groups in order to maintain a relatively constant γ_r across the series. In other words, it is unlikely that a QSAR developed for mononitro compounds will apply to dinitro or non-nitro compounds. The data in Fig. 1 and 3 support this position. Fig. 1 includes both mono- and dinitro compounds, and the structures in Fig. 3 contain 0-3 nitro groups. Much of the scatter in these Figures arises from variations in γ_r.

In summary, we have shown that inclusion of a term in octanol solubility may

improve a QSAR if the outliers have a physicochemical rather than a biological basis. We suggest that a K_{OW}-based QSAR constructed with groups of similar γ_r will be the most successful.

Acknowledgment

Portions of this work were used by CLW as partial fulfillment of the requirements for the Master of Science degree at the Institute of Paper Science and Technology.

REFERENCES

1. Deneer, J.W., T.L. Sinnige, W. Seinen and J.M. Hermens. 1987. Quantitative structure-activity relationships for the toxicity and bioconcentration factor of nitrobenzene derivatives towards the guppy (Poecilia reticulata). *Aquat. Toxicol.* 10:115-129.
2. Banerjee, S. and G.L. Baughman. 1991. Bioconcentration factors and lipid solubility. *Environ. Sci. Technol.* 25:536-539.
3. Purdy, R. 1988. Quantitative structure relationships for predicting toxicity of nitro-benzenes, phenols, anilines, and alkylamines to fathead minnows. *Proceedings*, 3rd. International Workshop on Quantitative Structure-Activity Relationships in Environmental Toxicology, Knoxville, TN, May 22-26, 1988, pp99-110.
4. Hansen, H.K., P. Rasmussen, A. Fredenslund, M. Schiller and J. Gmehling. 1991. Vapor-Liquid Equilibria by UNIFAC Group Contribution. 5. Revision and Extension. *Ind. Eng. Chem. Res.* 30:2352-2355.
5. Gmehling, J., P. Rasmussen and A. Fredenslund. 1981. Vapor-Liquid Equilibria by UNIFAC Group Contribution. 2. *Ind. Eng. Chem. Process Des. Dev.* 21:118-127.

TABLE 1: Values of γr at 25°C in octanol[1]

	γ_r		γ_r
-SiH3	0.049	-CH2SH	1.085
arom -C-OH	0.055	arom -CNH2	1.131
-CH2NH2 (amine)	0.124	-COOH	1.335
arom $-CH_2$	0.226	-CH=O	1.355
pyridine[2]	0.229	-OH	1.632
arom -C-F	0.324	thiophene[2]	1.699
-CH3	0.383	-I	1.751
arom -C-CH3	0.481	CH3C=O (ketone)	1.905
-(CONH2) amide	0.492	-CH2CN (nitrile)	1.908
arom -C-Cl	0.513	acetate	1.962
-CH2O- (ether)	0.520	ester (COO)	2.257
-CH=CH-	0.530	HCOO- (formate)	3.023
-CH2Cl	0.777	-CH2NO2	5.129
-Br	0.932	arom -CNO2	5.921
-CH2S- (sulfide)	0.958	water	7.422

[1]calculated using the parameter in ref. (5) except for arom-C-Cl, where the parameters in ref. (4) were used

[2]includes other pyridine- or thiophene-based heterocycles

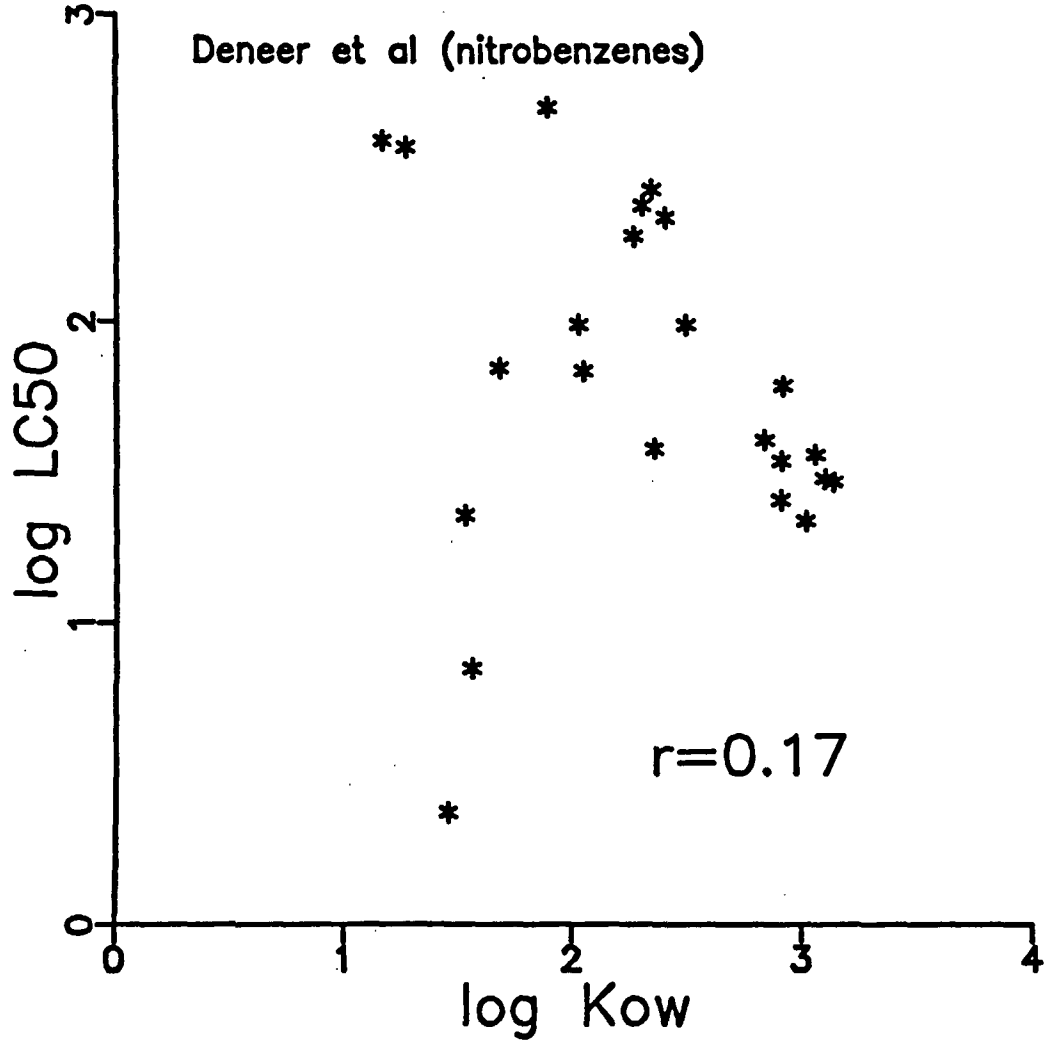


Fig.1 Log LC_{50}-log K_{ow} profiles of nitroaromatics containing up to two nitro-, chloro-, methyl or amino groups towards the guppy.

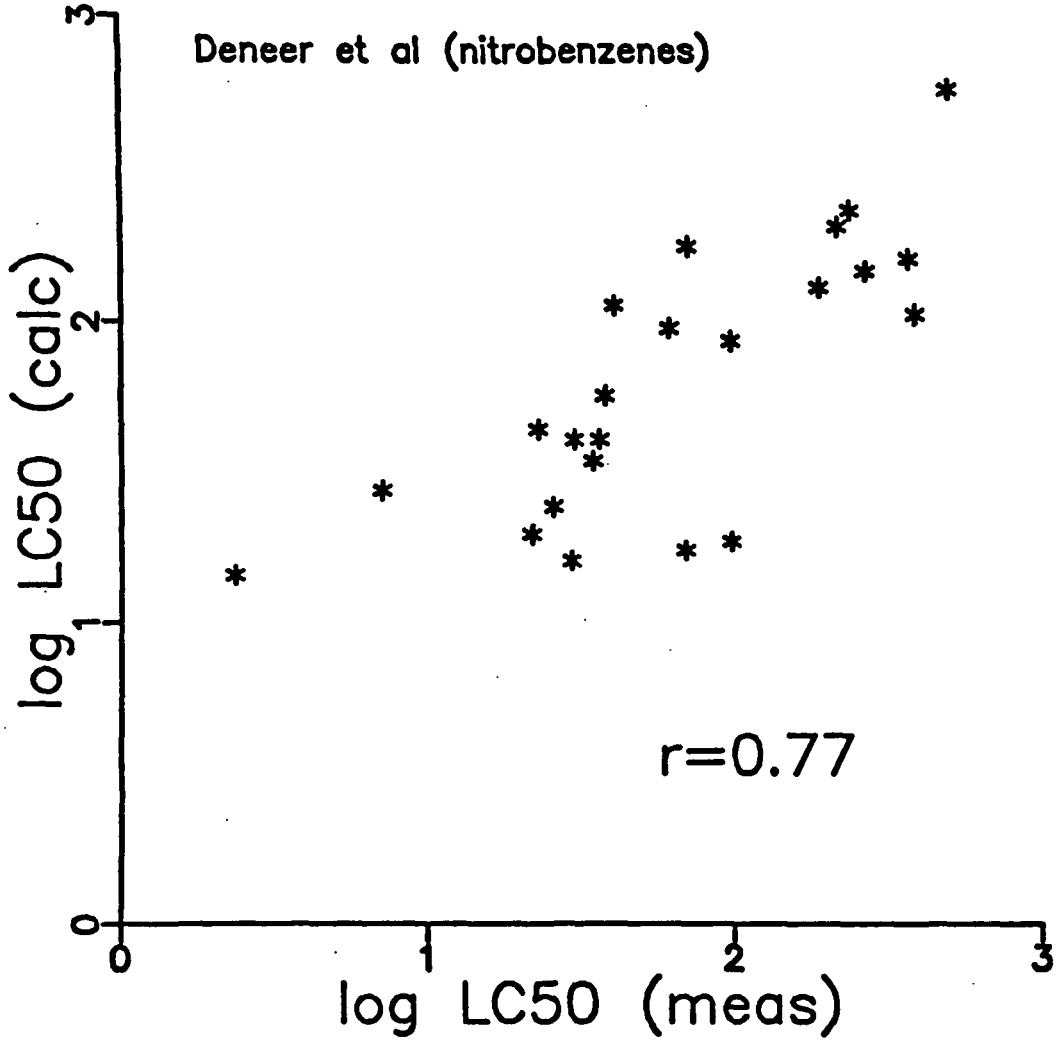


Fig. 2 Influence of the S_{oct} term on the Fig. 1 relationship. The calculated values are from eq 1.

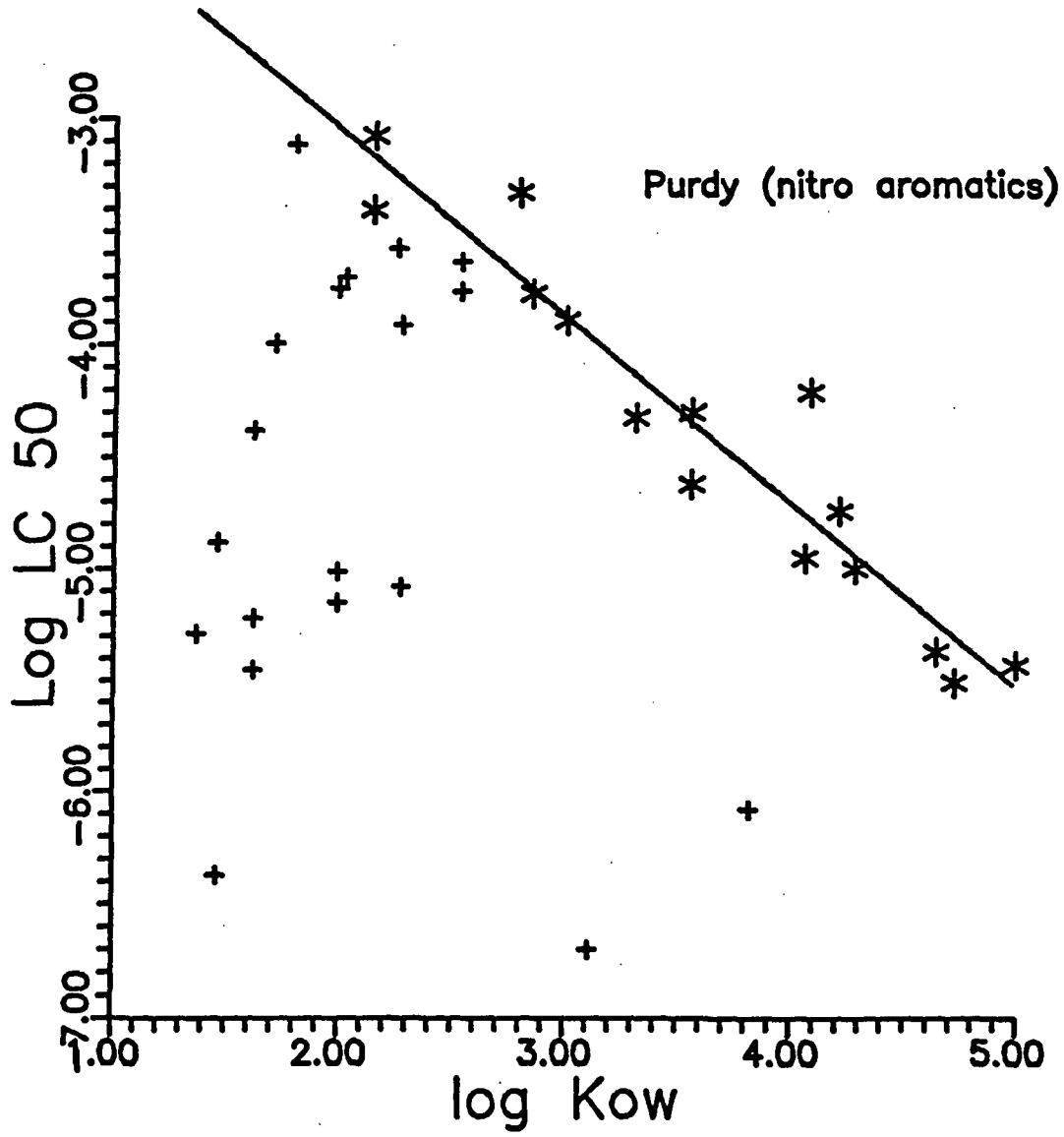


Fig. 3 Log LC_{50}-log K_{ow} profiles of aromatics containing up to three nitro groups, four chloro groups and other functionalities. The line is drawn through non nitro-compounds.

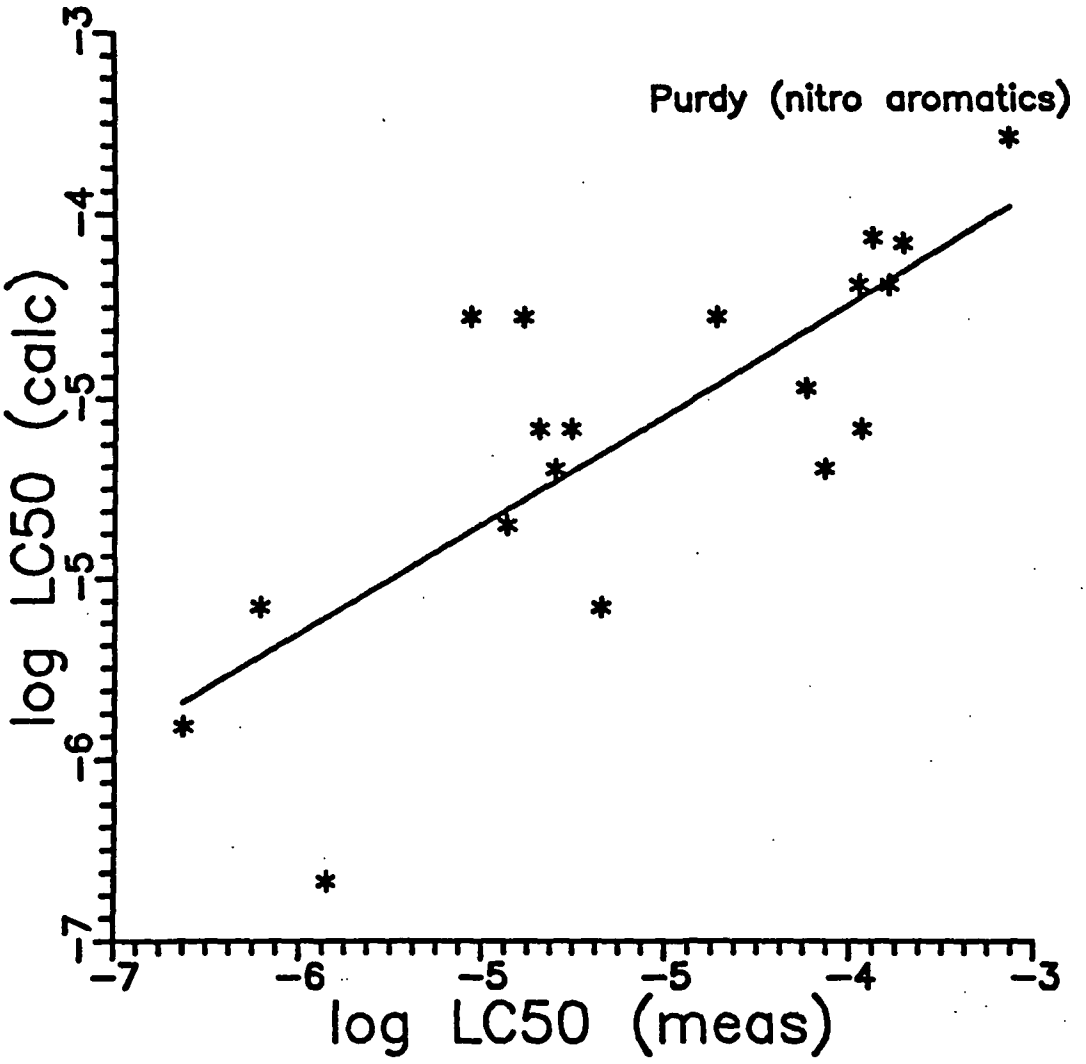


Fig. 4. Influence of the S_{oct} term on the Fig. 3 relationship. The calculated values are from eq 2.